AHRQ Quality Indicators
Patient Safety Indicators (PSI)
Composite Measure

1. Introduction

Many users of the AHRQ Quality Indicators (AHRQ QI) have expressed interest in the development of one or more composite measures. In particular, the National Healthcare Quality Report and the National Healthcare Disparities Report[1] asked the AHRQ QI program to develop composite measures for use in these reports. A composite measure for the Prevention Quality Indicators was developed initially[2]. The goal of the development effort is to develop a composite measure that might be used to monitor performance over time or across regions and populations using a methodology that applied at the national, regional, state or provider/area level. This report describes the construction of a composite measure for the Patient Safety Indicators: *Patient Safety for Selected Indicators*.

To assist in the development of a composite measure methodology, the AHRQ QI Composite Measure Workgroup held several conference calls to discuss important issues and considerations, and to provide feedback on preliminary results. In order to maintain the focus on the general composite measure methodology, the Workgroup did not consider the merits of including individual indicators in the composites. Rather, all available Patient Safety Indicators that met the conceptual criteria were included. The members of the AHRQ QI Composite Measure Workgroup are listed in Appendix A.

2. Why Composite Measures

Before considering alternative approaches to composite measures, one might consider why composite measures are potentially useful and for what purpose.

2.1 Benefits of Composite Measures

Composite measures have several potential benefits over individual indicators.

Summarize quality across multiple indicators. There are twenty (20) provider-level Patient Safety Indicators for various types of adverse events, making it difficult to formulate general statements about overall trends or differences in quality and patient safety.

Improve ability to detect quality differences. Combining information from multiple indicators may result in greater discrimination in performance than is evident from individual indicators.

[1] The most recent National Healthcare Quality Report and National Healthcare Disparities Report may be found at http://qualitytools.ahrq.gov.
[2] A report describing the composite measure for the Prevention Quality Indicators can be found at: http://www.qualityindicators.ahrq.gov/downloads/technical/AHRQ QI PQI Composite Report Final.pdf.

Identify important domains and drivers of quality. To the extent that certain indicators track together, or track with certain process or structure characteristics of providers, one may identify the important domains and drivers of quality.

Prioritize action for quality improvement. Individual indicators that contribute a larger share to the composite may be targets for quality improvement activity.

Make current decisions about future (unknown) healthcare needs. Depending on how the component indicators are weighted, composites may reflect the likely health outcomes for an individual or population.

Avoid cognitive "short-cuts". Research suggests that individuals faced with too many factors in making a decision take cognitive short-cuts that might not be in their best interest. Composites may help to ensure that decisions are made appropriately.

2.2 Concerns about Composite Measures.

Despite these benefits, there are concerns about using composite measures, depending on how the composite measure is constructed.

Mask important differences and relationships among components. Composite measures might mask the fact that two components are inversely related, or an "average" provider might be high on one component and low on another.

Not actionable. It might not be clear what action a provider should take given high or low performance on a composite measure.

Identify which parts of the healthcare system contribute most to quality. To the extent that the composite is not connected to the interventions important for the component measures, it might be difficult to know how the composite contributes to improving patient safety.

Detract from the impact and credibility of reports. The composite measure might not reflect the evidence-base of the individual indicators.

2.3 Potential Uses of Composite Measures.

Composite measures have many potential uses.

Consumers. Consumers might use composite measures to select a hospital or health plan either before or after a health event.

Providers. Providers might use composite measures to identify the domains and drivers of quality.

Purchasers. Purchasers might use composite measures to select hospitals or health plans in order to improve the health of employees.

Policymakers. Policymakers might use composite measures to set policy priorities in order to improve the health of a population

3. Alternative Perspectives on Composite Measures

There are two alternative perspectives on composite measures that guide the development of a composite measure methodology.

Signaling perspective. The signaling perspective seeks to guide decision making by providing information that will result in actions leading to some intended result. The ultimate evaluation criterion for the composite measure is the usefulness of the measure for achieving the intended result. An example of a composite measure reflecting the signaling perspective is the Dow Jones Industrial Average used to guide decision making on allocating investment resources.

Psychometric perspective. The psychometric perspective seeks to capture an underlying construct of quality based on multiple single indicators. The ultimate evaluation criterion for the composite measure is the extent to which the individual components reflect that construct. An example of a composite measure reflecting the psychometric perspective is the IQ test used to capture a construct labeled "intelligence."

The methodology used for the AHRQ QI composite measures reflects the signaling perspective, in that the primary intent of the measures is to guide decision making in terms of where to allocate resources to improve quality rather than to capture an underlying construct of quality.

4. Methodology for the AHRQ QI Composite Measures

4.1 Composite Measure Development Criteria

This report describes the construction of a single composite measure for the Patient Safety Indicators: *Patient Safety for Selected Indicators*. The basic criteria used to guide the development of the methodology were:

- *Evidence-based.* The composite measure should be based on indicator components that are important, reliable, valid, and minimally biased.
- *Conceptually coherent.* The components of the composite measure should be related to one another conceptually.
- *Empirically coherent.* The components of the composite measure should be related to one another empirically.
- *Intended use.* The composite measures should be constructed in a manner appropriate to the intended use, whether that is comparative reporting or quality improvement.

Applying these criteria to the Patient Safety Indicators, one could advocate for separate composites based on the type of adverse event (e.g., postoperative, etc.). However, in general the individual indicators apply to the same providers (see Table 2) and are positively correlated with each other, although not strongly so (see Table 3). Therefore the initial composite includes all the provider level (non-OB) indicators. Future development might examine sub-composites for certain indicators.

AHRQ PSI Composite Measure

Patient Safety for Selected Indicators	
PSI #03 Decubitus Ulcer	PSI #10 Postop Physio Metabol Derangmt
PSI #04 Failure To Rescue	PSI #11 Postop Respiratory Failure
PSI #06 Iatrogenic Pneumothorax	PSI #12 Postop PE Or DVT
PSI #07 Infection Due To Medical Care	PSI #13 Postop Sepsis
PSI #08 Postop Hip Fracture	PSI #14 Postop Wound Dehiscence
PSI #09 Postop Hemorrhage or Hematoma	PSI #15 Accidental Puncture/Laceration

4.2 The AHRQ QI Composite Measure Methodology

The general methodology for the AHRQ QI composite measures might be described as constructing a “composite of composites.” The first “composite” is the reliability-adjusted rate, which is a weighted average of the risk-adjusted rate and the reference population rate, where the weight is determined empirically. The second “composite” is a weighted average of the component indicators, where the weights are selected based on the intended use of the composite measure. These weights might be determined empirically or based on non-empirical considerations.

4.3 Constructing the AHRQ QI Composite Measure

The basic steps for computing the composite are as follows:

Step 1. Compute the risk-adjusted rate and confidence interval.

The AHRQ QI risk-adjusted rate is computed based on a logistic regression model[3] for calculating a predicted value for each case, and then summing the predicted value among all the cases in the hospital to compute the expected rate. The risk-adjusted rate is computed using indirect standardization as the observed rate divided by the expected rate multiplied by the reference population rate. The current reference population is the states participating in the HCUP program for 2001-2003, consisting of approximately 38 states and 90 million discharges[4].

[3] A separate workgroup is evaluating alternative risk-adjustment and hierarchical modeling methodologies for the AHRQ QI.

[4] The state data organizations that participated in the 2001-03 HCUP SID: Arizona Department of Health Services; California Office of Statewide Health Planning & Development; Colorado Health & Hospital Association; Connecticut - Chime, Inc.; Florida Agency for Health Care Administration; Georgia: An Association of Hospitals & Health Systems; Hawaii Health Information Corporation; Illinois Health Care Cost Containment Council; Indiana Hospital & Health Association; Iowa Hospital Association; Kansas Hospital Association; Kentucky Department for Public Health; Maine Health Data Organization; Maryland Health Services Cost Review; Massachusetts Division of Health Care Finance and Policy; Michigan Health & Hospital Association; Minnesota Hospital Association; Missouri Hospital Industry Data Institute; Nebraska Hospital Association; Nevada Department of Human Resources; New Hampshire Department of Health & Human Services; New Jersey Department of Health & Senior

Step 2. Scale the risk-adjusted rate using the reference population.

Table 1 shows the reference population numerator, denominator and rate for each of the Patient Safety Indicators. The levels of the rates vary from indicator to indicator. In order to combine the component indicators using a common scale, each indicator is first divided by the reference population rate. The components of the composite are therefore defined as deviations (i.e. a ratio) from the overall mean for each indicator.

Step3. Compute the reliability-adjusted rate.

The reliability-adjusted rate is computed as the weighted average of the risk-adjusted rate and the reference population rate, where the weights vary from zero to one, depending on degree of reliability for the indicator and provider (or other unit of analysis).

reliability-adjusted rate = [risk-adjusted rate * weight] + [reference population rate * (1 – weight)]

Table 4 shows the average reliability weights for the Patient Safety Indicators based on denominator size. For small providers, the weight is closer to zero. For large providers, the weight is closer to one. For a given provider, if the denominator is zero, then the weight assigned is zero (i.e., the reliability-adjusted rate is the reference population rate).

Step 4. Select the component weights

The composite measure is the weighted average of the scaled and reliability-adjusted rates for the component indicators. Table 5 shows examples of alternative weights that might be used. Other weights are also possible.

Single indicator weight. In this case, the composite is simply the reliability-adjusted rate for a single indicator.

Equal weight. In this case, each component indicator is assigned an identical weight based on the number of indicators. That is, the weight is equal to one divided by the number of indicators in the composite.

Services; New York State Department of Health; North Carolina Department of Health and Human Services; Ohio Hospital Association; Oregon Association of Hospitals & Health Systems; Pennsylvania Health Care Cost Containment Council; Rhode Island Department of Health; South Carolina State Budget & Control Board; South Dakota Association of Healthcare Organizations; Tennessee Hospital Association; Texas Health Care Information Council; Utah Department of Health; Vermont Association of Hospitals and Health Systems; Virginia Health Information; Washington State Department of Health; West Virginia Health Care Authority; Wisconsin Department of Health & Family Services.

Numerator weight. A numerator weight is based on the relative frequency of the numerator for each component indicator in the reference population. In general, a numerator weight reflects the amount of harm in the outcome of interest, in this case a potentially preventable adverse event. One might also use weights that reflect the amount of excess mortality or complications associated with the adverse event, or the amount of confidence one has in identifying events (i.e. the positive predictive value).

Denominator weight. A denominator weight is based on the relative frequency of the denominator for each component indicator in the reference population. In general, a denominator weight reflects the amount of risk for experiencing the outcome of interest for a given population. For example, the denominator weight might be based on the demographic composition of a health plan, the employees of a purchaser, a state, an individual hospital, or a single individual patient.

Factor weight. A factor weight is based on some sort of analysis which assigns each component indicator a weight that reflects the contribution of that indicator to the common variation among the indicators. The component indicator that is most predictive of that common variation is assigned the highest weight. The weights in Table 5 are based on a principal components factor analysis of the reliability-adjusted rates.

Step 5. Construct the composite measure

The composite measure is the weighted average of the component indicators using the selected weights and the scaled and reliability-adjusted indicators.

Composite = [indicator1 * weight1] + [indicator2 * weight2] + . . . + [indicatorN * weightN]

The confidence interval on the composite is based on the standard error of the composite, which is the square root of the variance. The variance is computed based on the signal variance-covariance matrix and the reliability weights. Details of the computation are provided in the appendix.

4.4 An example computation of the composite measure

This example demonstrates the construction of the composite for a representative provider beginning with the risk-adjusted rate and standard error for each Patient Safety Indicator. An important consideration in the development of the composite measure methodology was that the computation of the composite and the weights used be transparent and that a provider should be able to trace the computation from the individual indicators to the composite and back again.

Step 1. Compute the risk-adjusted rate and confidence interval

Table S1. A Single Provider in a Single Year

PSI	Patients	Observed Rate	Risk-adjusted Rate	Risk-adjusted SE
PSI #03 Decubitus Ulcer	4,174	15.78	30.18	3.08
PSI #04 Failure To Rescue	832	97.41	110.97	12.14
PSI #06 Iatrogenic Pneumothorax	14,454	0.90	1.00	0.20
PSI #07 Infection Due To Medical Care	11,031	3.59	3.13	0.39
PSI #08 Postop Hip Fracture	2,174	0.41	0.30	0.28
PSI #09 Postop Hemorrhage or Hematoma	4,529	3.23	3.03	0.68
PSI #10 Postop Physio Metabol Derangmt	2,521	0.87	0.55	0.52
PSI #11 Postop Respiratory Failure	2,130	9.30	7.22	1.78
PSI #12 Postop PE Or DVT	4,508	18.65	15.86	1.33
PSI #13 Postop Sepsis	541	9.07	7.94	3.90
PSI #14 Postop Wound Dehiscence	727	0.00	0.00	1.67
PSI #15 Accidental Puncture/Laceration	16,174	6.17	9.56	0.58

Note: Observed and risk-adjusted rate are per 1,000

This is the output a user would obtain from applying the AHRQ QI software (SAS and Windows) to the user's data.

Step 2. Scale the risk-adjusted rate using the reference population

Table S2. Scaling the Single Provider Rate

PSI	Reference Population Rate	Risk-adjusted Ratio	Risk-adjusted SE
PSI #03 Decubitus Ulcer	22.08	1.3670	0.1396
PSI #04 Failure To Rescue	135.62	0.8182	0.0895
PSI #06 Iatrogenic Pneumothorax	0.58	1.7219	0.3502
PSI #07 Infection Due To Medical Care	2.05	1.5288	0.1900
PSI #08 Postop Hip Fracture	0.27	1.0960	1.0285
PSI #09 Postop Hemorrhage or Hematoma	2.17	1.3955	0.3149
PSI #10 Postop Physio Metabol Derangmt	0.99	0.5544	0.5225
PSI #11 Postop Respiratory Failure	8.80	0.8208	0.2022
PSI #12 Postop PE Or DVT	9.17	1.7293	0.1448
PSI #13 Postop Sepsis	10.01	0.7932	0.3899
PSI #14 Postop Wound Dehiscence	2.08	0.0000	0.8023
PSI #15 Accidental Puncture/Laceration	3.61	2.6485	0.1605

The individual indicators are scaled by the reference population rate so that each indicator reflects the degree of deviation from the overall average performance.

Step 3. Compute the reliability-adjusted rate

Table S3A. Compute the Reliability Weight

PSI	Risk-adjusted SE	Noise Variance	Signal Variance	Reliability Weight
PSI #03 Decubitus Ulcer	0.1396	0.0195	0.2369	0.9240
PSI #04 Failure To Rescue	0.0895	0.0080	0.0408	0.8357
PSI #06 Iatrogenic Pneumothorax	0.3502	0.1226	0.1819	0.5974
PSI #07 Infection Due To Medical Care	0.1900	0.0361	0.2974	0.8917
PSI #08 Postop Hip Fracture	1.0285	1.0577	1.0925	0.5081
PSI #09 Postop Hemorrhage or Hematoma	0.3149	0.0992	0.1234	0.5544
PSI #10 Postop Physio Metabol Derangmt	0.5225	0.2730	0.2538	0.4818
PSI #11 Postop Respiratory Failure	0.2022	0.0409	0.2733	0.8699
PSI #12 Postop PE Or DVT	0.1448	0.0210	0.2047	0.9071
PSI #13 Postop Sepsis	0.3899	0.1520	0.1797	0.5418
PSI #14 Postop Wound Dehiscence	0.8023	0.6437	0.2676	0.2937
PSI #15 Accidental Puncture/Laceration	0.1605	0.0258	0.2168	0.8938

Note: Noise variance is standard error squared; Reliability weight is signal variance / (signal variance + noise variance)

The noise variance is computed from the user's data as the square root of the standard error. The signal variance is a reference population parameter that reflects the amount of provider level variation remaining after the noise variance is removed. Note that the noise variance will vary by provider and by indicator.

Table S3B. Compute the Reliability-adjusted Rate

PSI	Reliability Weight	Risk-adjusted Ratio	Reference Population Ratio	Reliability-adjusted Ratio
PSI #03 Decubitus Ulcer	0.9240	1.3670	1.0000	1.3391
PSI #04 Failure To Rescue	0.8357	0.8182	1.0000	0.8481
PSI #06 Iatrogenic Pneumothorax	0.5974	1.7219	1.0000	1.4313
PSI #07 Infection Due To Medical Care	0.8917	1.5288	1.0000	1.4715
PSI #08 Postop Hip Fracture	0.5081	1.0960	1.0000	1.0488
PSI #09 Postop Hemorrhage or Hematoma	0.5544	1.3955	1.0000	1.2193
PSI #10 Postop Physio Metabol Derangmt	0.4818	0.5544	1.0000	0.7853
PSI #11 Postop Respiratory Failure	0.8699	0.8208	1.0000	0.8441
PSI #12 Postop PE Or DVT	0.9071	1.7293	1.0000	1.6615
PSI #13 Postop Sepsis	0.5418	0.7932	1.0000	0.8880
PSI #14 Postop Wound Dehiscence	0.2937	0.0000	1.0000	0.7063
PSI #15 Accidental Puncture/Laceration	0.8938	2.6485	1.0000	2.4734

Note Reliability-adjusted ratio is [risk-adjusted ratio * weight] + [reference population ratio * (1 – weight)]

The first "composite" is the weighted average of the provider's risk-adjusted ratio and the reference population ratio, where the weight reflects the reliability of the provider's risk-adjusted ratio.

Step 4. Select the component weights

Table S4. Denominator Weight

	Denominator Weight
PSI #03 Decubitus Ulcer	0.0749
PSI #04 Failure To Rescue	0.0086
PSI #06 Iatrogenic Pneumothorax	0.2226
PSI #07 Infection Due To Medical Care	0.1848
PSI #08 Postop Hip Fracture	0.0469
PSI #09 Postop Hemorrhage or Hematoma	0.0706
PSI #10 Postop Physio Metabol Derangmt	0.0341
PSI #11 Postop Respiratory Failure	0.0278
PSI #12 Postop PE Or DVT	0.0703
PSI #13 Postop Sepsis	0.0085
PSI #14 Postop Wound Dehiscence	0.0151
PSI #15 Accidental Puncture/Laceration	0.2357

The weights are selected depending on the intended use of the composite. In this example, we use the denominator weight.

Step 5. Construct the composite measure

Table S5. Construct the Composite Measure

	Denominator Weight (A)	**Reliability-adjusted Ratio (B)**	**(A) *(B)**
PSI #03 Decubitus Ulcer	0.0749	1.3391	0.1003
PSI #04 Failure To Rescue	0.0086	0.8481	0.0073
PSI #06 Iatrogenic Pneumothorax	0.2226	1.4313	0.3186
PSI #07 Infection Due To Medical Care	0.1848	1.4715	0.2719
PSI #08 Postop Hip Fracture	0.0469	1.0488	0.0492
PSI #09 Postop Hemorrhage or Hematoma	0.0706	1.2193	0.0861
PSI #10 Postop Physio Metabol Derangmt	0.0341	0.7853	0.0268
PSI #11 Postop Respiratory Failure	0.0278	0.8441	0.0235
PSI #12 Postop PE Or DVT	0.0703	1.6615	0.1168
PSI #13 Postop Sepsis	0.0085	0.8880	0.0075
PSI #14 Postop Wound Dehiscence	0.0151	0.7063	0.0107
PSI #15 Accidental Puncture/Laceration	0.2357	2.4734	0.5830

Patient Safety for Selected Indicators		*1.6016*
Standard Error		*0.0959*
Confidence Interval at $p<0.05$	*1.7896*	*1.4137*

The final composite is simply the weighted average of the component indicators. Note the potential application of the composite construction for use in quality improvement. The final computation shows that Accidental Puncture/Laceration is the largest single contributor to the composite both because the indicator was heavily weighted and because the performance of the provider was worse than average. The incentive created in using the composite is to allocate resources to reducing Accidental Puncture/Laceration as the best mechanism to lowering the composite score.

5. Performance of the AHRQ QI Composite Measures

5.1 Evaluation Criteria

The tables and figures show the performance of each composite measure. The composite measures are evaluated using three criteria: discrimination, forecasting and construct validity.

Discrimination is the ability of the composite measure to differentiate performance as measured by statistically significant deviations from the average performance.

Forecasting is the ability of the composite measure to predict performance for each of the component indicators. Ideally, the forecasting performance would reflect the weighting of the components, in the sense that forecasting would maximize the differences for the most highly weighted components.

Construct validity is the degree of association between the composite and other aggregate measures of quality. In this report we look primarily at the consistency in the composite over time. A broader analysis of construct validity would examine the relationship between the composites and external measures of quality and patient safety or other factors that might influence quality and patient safety.

5.2 Results

Table 6 shows the discrimination performance of the composite measure *Patient Safety for Selected Indicators*. The columns show the percent of providers that are either worse than average or better than average based on the confidence interval for the composite measure. The discrimination performance varies depending on the weight used. The equal weight has the least ability to discriminate. The single indicator used as an example is “selected infection due to medical care”. The numerator weight tends to have the greatest ability to discriminate, followed by the denominator weight, the factor weight and equal weight.

In general, the composite identifies a large number of providers with performance that is better or worse than average. Figure 4 shows the distribution of each composite and the 95 percent confidence interval.

Table 7 shows the forecasting performance of the composite measure. In this analysis each provider is assigned to a quintile (Q1-Q5) based on the performance on the composite in 2001-2002. The columns show the relative difference in the actual risk-adjusted rate in 2003 for the best and worst performing quintile relative to the middle sixty percent.

Forecasting performance varies depending on the weights used to construct the composite. In general, the composite is better at forecasting performance on component indicators that are more heavily weighted.

Table 8 shows the correlation among the composite measures using the alternative weights. For the *Patient Safety for Selected Indicators*, the correlations vary from 0.70 to 0.90.

The table also shows the correlation in the composite measures from one year to the next. For the *Patient Safety for Selected Indicators*, the correlation depends on the weight used, with the numerator and factor weight showing the most persistence.

6. Concluding Comments

The intent of the AHRQ QI Composite Measure project was to develop a general methodology that could be used primarily to monitor performance in national and regional reporting, but that also could be applied to comparative reporting and quality improvement at the provider level. An important caveat in using the composite measures is that the measures are not intended to reflect any broader construct of quality or patient safety than is reflect in the individual indicators themselves, and that the composites are only has useful and valid as are the individual indicators that make up the composite. As the AHRQ Quality Indicators and the data upon which they are based continue to improve, the composite measures will improve as potentially useful tools for decision making in allocating quality improvement resources.

Appendix A. AHRQ QI Composite Measure Workgroup

Workgroup Members

- John Birkmeyer, University of Michigan
- Bruce Boissonnault, Niagara Health Quality Coalition
- John Bott, Employer Health Care Alliance Cooperative
- Dale Bratzler, Oklahoma Foundation for Medical Quality
- Sharon Cheng, MedPAC
- Elizabeth Clough, Wisconsin Collaborative for Healthcare Quality
- Nancy Dunton, University of Kansas Medical Center, School of Nursing
- John Hoerner, Hospital Industry Data Institute
- David Hopkins, Pacific Business Group on Health
- Gregg Meyer, Massachusetts General Physicians Organization
- Elizabeth Mort, Massachusetts General
- Janet Muri, National Perinatal Information Center
- Vi Naylor, Georgia Hospital Association
- Eric Peterson, Duke University Medical Center
- Martha Radford, New York University Hospitals Center
- Gulzar Shah, National Association of Health Data Organizations
- Paul Turner, Vermont Program for Quality in Health Care

Liaison Members

- Justine Carr, National Committee on Vital and Health Statistics
- Robert Hungate, National Committee on Vital and Health Statistics
- Sheila Roman, Centers for Medicare & Medicaid Services
- Amy Rosen, Bedford Veterans Affairs Medical Center
- Stephen Schmaltz, Joint Commission on Accreditation of Healthcare Organizations
- Jane Sisk, National Center for Health Statistics
- Ernie Moy, Agency for Healthcare Research and Quality

Technical Advisors

- John Adams, RAND
- Bob Houchens, Medstat
- Bill Rogers, Rogers Associate
- Chunliu Zhan, Agency for Healthcare Research and Quality

AHRQI QI Support

- Mamatha Pancholi, AHRQ QI Project Officer
- Marybeth Farquhar, AHRQ NQF Project Officer
- Jeffrey Geppert, Project Director, Battelle Memorial Institute
- Theresa Schaaf, Project Manager, Battelle Memorial Institute
- Douglas O. Staiger, Technical Consultant, Dartmouth College

Appendix B. PSI Composite Tables

Table 1. Reference Population

PSI	Numerator	Denominator	Rate
PSI #03 Decubitus Ulcer	476,583	21,583,071	22.08
PSI #04 Failure To Rescue	337,421	2,488,029	135.62
PSI #06 Iatrogenic Pneumothorax	37,335	64,193,131	0.58
PSI #07 Infection Due To Medical Care	109,442	53,292,737	2.05
PSI #08 Postop Hip Fracture	3,685	13,533,878	0.27
PSI #09 Postop Hemorrhage or Hematoma	44,250	20,347,679	2.17
PSI #10 Postop Physio Metabol Derangmt	9,700	9,841,216	0.99
PSI #11 Postop Respiratory Failure	70,440	8,002,305	8.80
PSI #12 Postop PE Or DVT	185,794	20,263,685	9.17
PSI #13 Postop Sepsis	24,633	2,461,073	10.01
PSI #14 Postop Wound Dehiscence	9,038	4,346,106	2.08
PSI #15 Accidental Puncture/Laceration	245,532	67,971,505	3.61

Source: HCUP State Inpatient Data, 2001-2003; Rate per 1,000

Table 2. Provider-level Rates

		Risk adjusted		Reliability adjusted	
PSI	**Providers**	**Rate**	**Std. Dev.**	**Rate**	**Std. Dev.**
PSI #03 Decubitus Ulcer	4,823	19.85	19.77	20.34	10.07
PSI #04 Failure To Rescue	4,476	117.98	70.63	133.10	21.57
PSI #06 Iatrogenic Pneumothorax	4,909	0.41	0.63	0.54	0.16
PSI #07 Infection Due To Medical Care	4,908	1.52	2.55	1.76	1.20
PSI #08 Postop Hip Fracture	4,312	0.42	8.23	0.28	0.12
PSI #09 Postop Hemorrhage or Hematoma	4,356	1.86	3.69	2.14	0.46
PSI #10 Postop Physio Metabol Derangmt	3,603	0.97	10.76	0.90	0.28
PSI #11 Postop Respiratory Failure	3,592	7.67	11.25	8.37	3.93
PSI #12 Postop PE Or DVT	4,352	7.85	12.06	8.29	3.21
PSI #13 Postop Sepsis	3,398	9.73	14.25	9.61	3.14
PSI #14 Postop Wound Dehiscence	4,004	1.98	3.38	2.10	0.59
PSI #15 Accidental Puncture/Laceration	4,909	2.72	3.37	3.21	1.35

Source: HCUP State Inpatient Data, 2001-2003; Rate per 1,000

Table 3. Provider-level Correlation

PSI	PSI #03	PSI #04	PSI #06	PSI #07	PSI #08	PSI #09	PSI #10	PSI #11
PSI #03 Decubitus Ulcer	1.000	0.230	0.106	0.198	0.034	-0.027	0.060	0.155
PSI #04 Failure To Rescue		1.000	0.134	0.122	-0.009	0.030	0.019	0.060
PSI #06 Iatrogenic Pneumothorax			1.000	0.404	0.003	0.194	0.042	0.056
PSI #07 Infection Due To Medical Care				1.000	0.005	0.174	0.082	0.116
PSI #08 Postop Hip Fracture					1.000	-0.005	0.000	-0.001
PSI #09 Postop Hemorrhage or Hematoma						1.000	0.016	0.024
PSI #10 Postop Physio Metabol Derangmt							1.000	0.130
PSI #11 Postop Respiratory Failure								1.000
PSI #12 Postop PE Or DVT								
PSI #13 Postop Sepsis								
PSI #14 Postop Wound Dehiscence								
PSI #15 Accidental Puncture/Laceration								

Source: HCUP State Inpatient Data, 2001-2003; Rate per 1,000

Table 3. Provider-level Correlation (continued)

PSI	PSI #12	PSI #13	PSI #14	PSI #15
PSI #03 Decubitus Ulcer	0.192	0.171	0.054	-0.003
PSI #04 Failure To Rescue	-0.101	0.045	0.067	0.029
PSI #06 Iatrogenic Pneumothorax	0.122	0.033	0.085	0.376
PSI #07 Infection Due To Medical Care	0.253	0.160	0.005	0.301
PSI #08 Postop Hip Fracture	-0.001	0.016	0.013	0.002
PSI #09 Postop Hemorrhage or Hematoma	0.034	-0.010	0.070	0.273
PSI #10 Postop Physio Metabol Derangmt	0.081	0.055	0.035	0.073
PSI #11 Postop Respiratory Failure	0.122	0.221	0.054	-0.037
PSI #12 Postop PE Or DVT	1.000	0.107	0.010	0.052
PSI #13 Postop Sepsis		1.000	0.006	-0.061
PSI #14 Postop Wound Dehiscence			1.000	0.054
PSI #15 Accidental Puncture/Laceration				1.000

Source: HCUP State Inpatient Data, 2001-2003; Rate per 1,000

Table 4. Reliability Weight by Denominator Size

PSI	Providers	Q1	Q2	Q3	Q4
Average Annual Denominator Size (by quartile)					
PSI #03 Decubitus Ulcer	4,823	53.3	367.3	1,325.2	4,219.7
PSI #04 Failure To Rescue	4,476	6.5	38.6	151.4	544.6
PSI #06 Iatrogenic Pneumothorax	4,909	276.6	1,248.1	3,939.3	11,965.4
PSI #07 Infection Due To Medical Care	4,908	210.7	1,020.1	3,285.3	9,961.7
PSI #08 Postop Hip Fracture	4,312	25.3	207.9	774.8	3,176.8
PSI #09 Postop Hemorrhage or Hematoma	4,356	30.5	305.8	1,181.1	4,710.8
PSI #10 Postop Physio Metabol Derangmt	3,603	17.2	162.4	650.2	2,811.1
PSI #11 Postop Respiratory Failure	3,592	15.9	150.1	574.4	2,230.1
PSI #12 Postop PE Or DVT	4,352	30.7	306.9	1,177.4	4,693.2
PSI #13 Postop Sepsis	3,398	6.6	46.0	164.1	748.5
PSI #14 Postop Wound Dehiscence	4,004	18.1	112.9	324.0	992.3
PSI #15 Accidental Puncture/Laceration	4,909	285.9	1,284.4	4,093.2	12,791.5

PSI	Q1	Q2	Q3	Q4	Weighted Average
Average Reliability Weight					
PSI #03 Decubitus Ulcer	0.4192	0.8061	0.9293	0.9819	0.9606
PSI #04 Failure To Rescue	0.0858	0.3885	0.7464	0.9076	0.8604
PSI #06 Iatrogenic Pneumothorax	0.0709	0.2227	0.5083	0.7730	0.6900
PSI #07 Infection Due To Medical Care	0.2603	0.5762	0.8272	0.9390	0.8906
PSI #08 Postop Hip Fracture	0.0438	0.1498	0.3889	0.6838	0.6453
PSI #09 Postop Hemorrhage or Hematoma	0.0249	0.1796	0.4666	0.7532	0.7089
PSI #10 Postop Physio Metabol Derangmt	0.0207	0.0896	0.2799	0.6247	0.5993
PSI #11 Postop Respiratory Failure	0.1483	0.5041	0.7834	0.9247	0.8908
PSI #12 Postop PE Or DVT	0.1517	0.5805	0.8554	0.9538	0.9264
PSI #13 Postop Sepsis	0.0644	0.2377	0.4677	0.7552	0.7238
PSI #14 Postop Wound Dehiscence	0.0296	0.1607	0.3510	0.5867	0.5298
PSI #15 Accidental Puncture/Laceration	0.1860	0.5497	0.8516	0.9600	0.9077

Source: HCUP State Inpatient Data, 2001-2003

Table 5. Alternative Composite Weights

	Single Indicator Weight	Equal Weight	Numerator Weight	Denominator Weight	Factor Weight
PSI #03 Decubitus Ulcer	0.0000	0.0833	**0.3067**	0.0749	0.1020
PSI #04 Failure To Rescue	0.0000	0.0833	**0.2172**	0.0086	0.0430
PSI #06 Iatrogenic Pneumothorax	0.0000	0.0833	0.0240	**0.2226**	0.0927
PSI #07 Infection Due To Medical Care	**1.0000**	0.0833	0.0704	0.1848	0.1415
PSI #08 Postop Hip Fracture	0.0000	**0.0833**	0.0024	0.0469	0.0222
PSI #09 Postop Hemorrhage or Hematoma	0.0000	**0.0833**	0.0285	0.0706	0.0613
PSI #10 Postop Physio Metabol Derangmt	0.0000	0.0833	0.0062	0.0341	**0.1085**
PSI #11 Postop Respiratory Failure	0.0000	0.0833	0.0453	0.0278	**0.1088**
PSI #12 Postop PE Or DVT	0.0000	0.0833	**0.1196**	0.0703	0.1098
PSI #13 Postop Sepsis	0.0000	0.0833	0.0159	0.0085	**0.1192**
PSI #14 Postop Wound Dehiscence	0.0000	**0.0833**	0.0058	0.0151	0.0330
PSI #15 Accidental Puncture/Laceration	0.0000	0.0833	0.1580	**0.2357**	0.0582

Source: HCUP State Inpatient Data, 2001-2003. . For each indicator, the most highly weighted composite is in **bold**.

Table 6. Discrimination Performance of Alternative Composites

Composite	Providers	%Better than Average	%Worse than Average	% Total
Patient Safety for Selected Indicators				
Single Indicator Weight	4,908	17.83%	11.67%	29.50%
Equal Weight	4,898	5.43%	6.17%	11.60%
Numerator Weight	4,899	27.09%	13.84%	40.93%
Denominator Weight	4,910	15.09%	9.33%	24.42%
Factor Weight	4,901	10.26%	8.43%	18.69%

Source: HCUP State Inpatient Data, 2001-2003

Table 7. Forecast Performance of Alternative Composites

PSI	PSI #03	PSI #04	PSI #06	PSI #07	PSI #08	PSI #09	PSI #10	PSI #11
Patient Safety for Selected Indicators								
Single Indicator Weight								
Best 20%	-0.071*	0.011	-0.224*	**-0.469***	0.042	-0.095*	-0.236**	-0.145*
Worst 20%	0.145*	0.030*	0.392*	**0.835***	0.143	0.161*	0.203**	0.161*
Equal Weight								
Best 20%	-0.253*	-0.057*	-0.239*	-0.331*	**-0.048**	**-0.139***	-0.328*	-0.267*
Worst 20%	0.297*	0.045*	0.495*	0.593*	**0.108**	**0.216***	0.449*	0.279*
Numerator Weight								
Best 20%	**-0.373***	**-0.098***	-0.225*	-0.427*	-0.071	-0.158*	-0.274*	-0.164*
Worst 20%	**0.571***	**0.085***	0.372*	0.402*	0.076	0.105*	0.380*	0.227*
Denominator Weight								
Best 20%	-0.182*	-0.024**	**-0.339***	-0.432*	-0.063	-0.165*	-0.243*	-0.112**
Worst 20%	0.189*	0.016	**0.578***	0.667*	-0.001	0.226*	0.325*	0.205*
Factor Weight								
Best 20%	-0.305*	-0.055*	-0.253*	-0.401*	-0.035	-0.127*	**-0.334***	**-0.328***
Worst 20%	0.264*	0.014	0.457*	0.655*	0.044	0.178*	**0.452***	**0.320***

Source: HCUP State Inpatient Data, 2001-2003; *Significant at p<.05; ** Significant at p<.10. The forecast predicts performance in 2003 (rate) based on performance in 2001-2002 (by quintile) using five alternative measures composite weights. For each indicator, the most highly weighted composite is in **bold**.

Table 7 (continued). Forecast Performance of Alternative Composites

PSI	PSI #12	PSI #13	PSI #14	PSI #15
Patient Safety for Selected Indicators				
Single Indicator Weight				
Best 20%	-0.148*	-0.137*	-0.015	-0.173*
Worst 20%	0.277*	0.114*	0.034	0.221*
Equal Weight				
Best 20%	-0.192*	-0.257*	**-0.119***	-0.163*
Worst 20%	0.371*	0.189*	**0.233***	0.318*
Numerator Weight				
Best 20%	**-0.208***	-0.222*	0.008	-0.262*
Worst 20%	**0.387***	0.097**	0.145*	0.264*
Denominator Weight				
Best 20%	-0.155*	-0.148*	-0.103	**-0.342***
Worst 20%	0.305*	0.042	0.063	**0.494***
Factor Weight				
Best 20%	-0.214*	**-0.335***	-0.068	-0.147*
Worst 20%	0.423*	**0.188***	0.148*	0.260*

Source: HCUP State Inpatient Data, 2001-2003; *Significant at p<.05; ** Significant at p<.10. The forecast predicts performance in 2003 (rate) based on performance in 2001-2002 (by quintile) using five alternative measures composite weights. For each indicator, the most highly weighted composite is in **bold**.

Table 8. Correlation of Alternative Composites

Composite	Single Indicator Weight	Equal Weight	Numerator Weight	Denominator Weight	Factor Weight	Year-to-Year
Patient Safety for Selected Indicators						
Single Indicator Weight	1.000	0.630*	0.517*	0.755*	0.714*	0.738
Equal Weight		1.000	0.794*	0.851*	0.962*	0.735
Numerator Weight			1.000	0.726*	0.801*	0.790
Denominator Weight				1.000	0.831*	0.773
Factor Weight					1.000	0.795

Source: HCUP State Inpatient Data, 2001-2003; *Significant at p<.05; ** Significant at p<.10. Year-to-year correlation in alternative composites between 2002 and 2003.

Appendix C. PSI Composite Figures

Figure 1. Provider Level Rates
Selected Infection Due To Medical Care

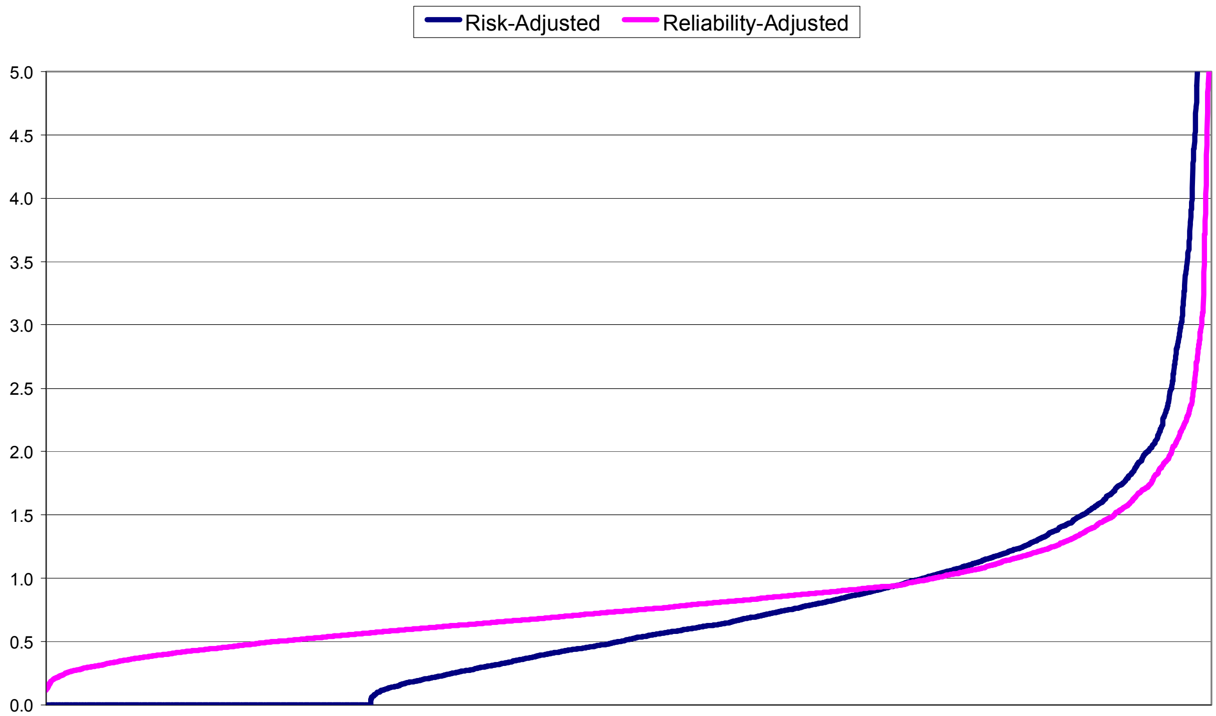

**Figure 2. Impact of Reliability Weight for a Single Provider
Selected Infection due to Medical Care**

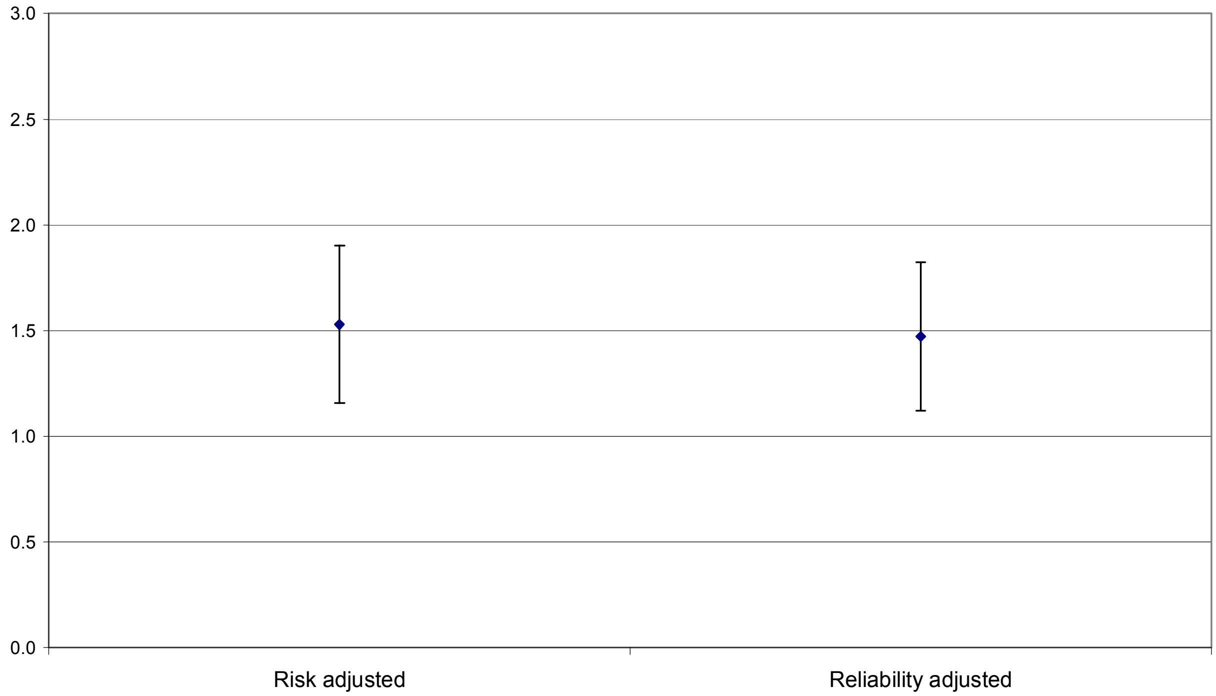

Figure 3. Impact of Reliablity Weight by Denominator Size
Selected Infection due to Medical Care

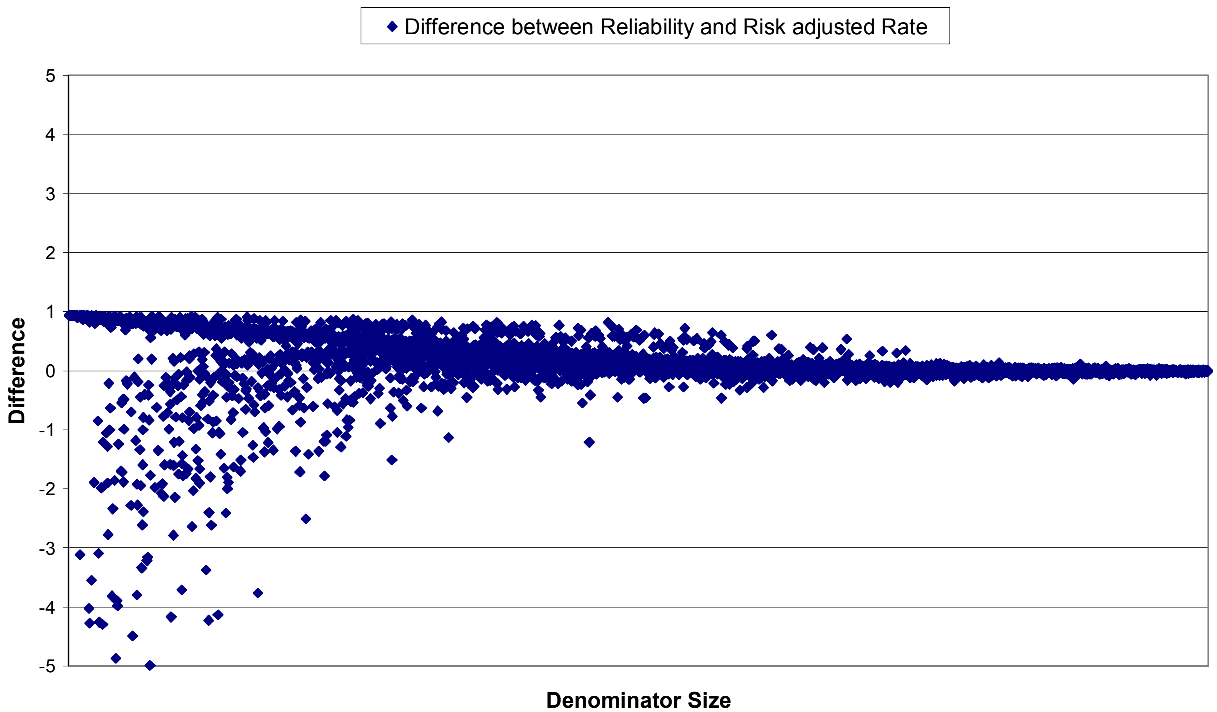

Figure 4-1. Patient Safety for Selected Indicators, Single Indicator Weight

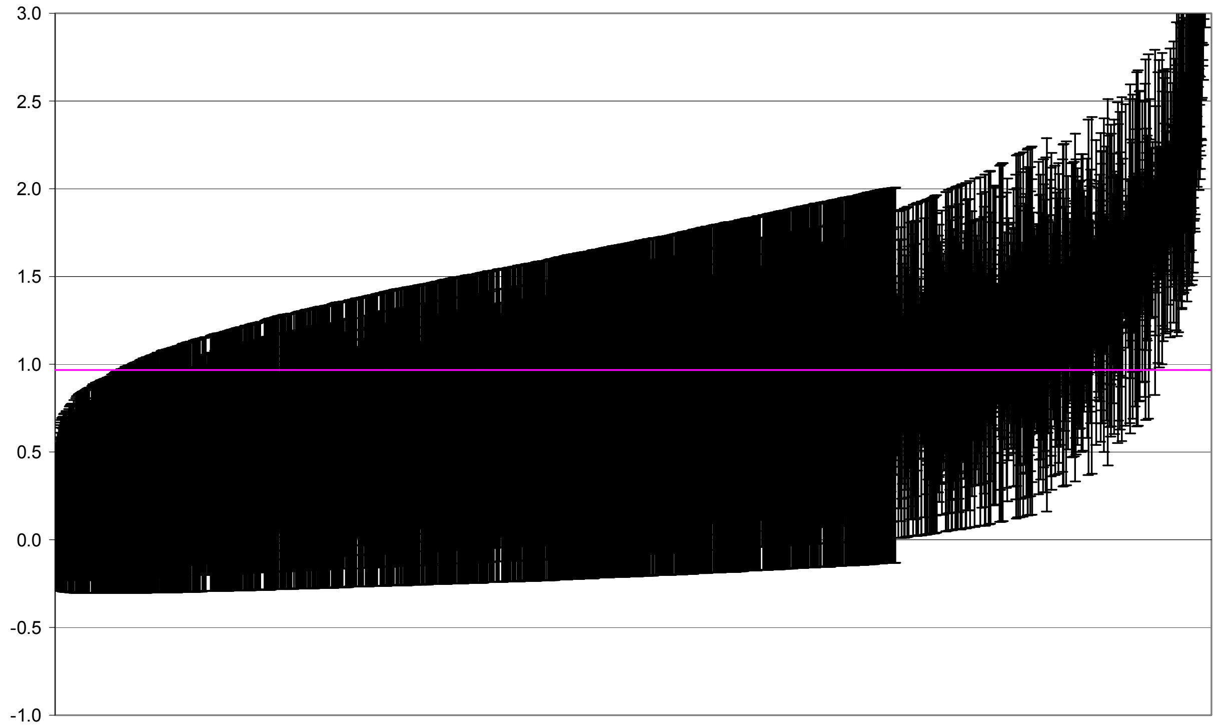

Figure 4-2. Patient Safety for Selected Indicators, Equal Weight

Figure 4-3. Patient Safety for Selected Indicators, Numerator Weight

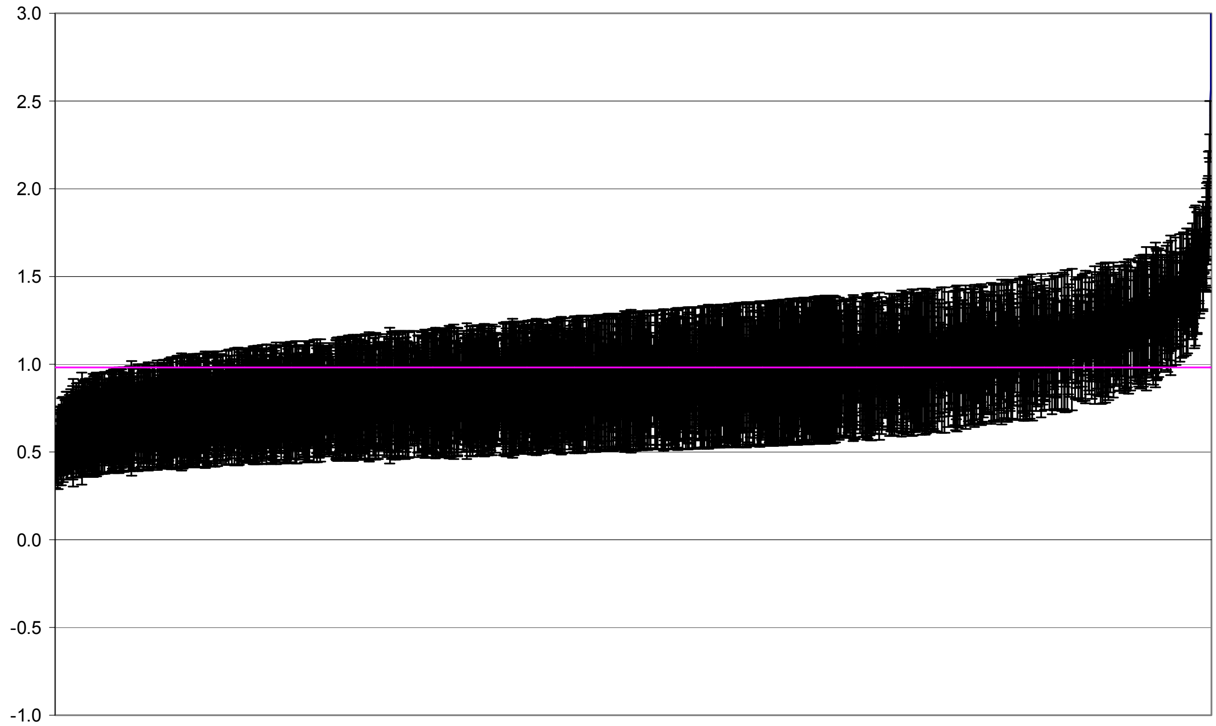

Figure 4-4. Patient Safety for Selected Indicators, Denominator Weight

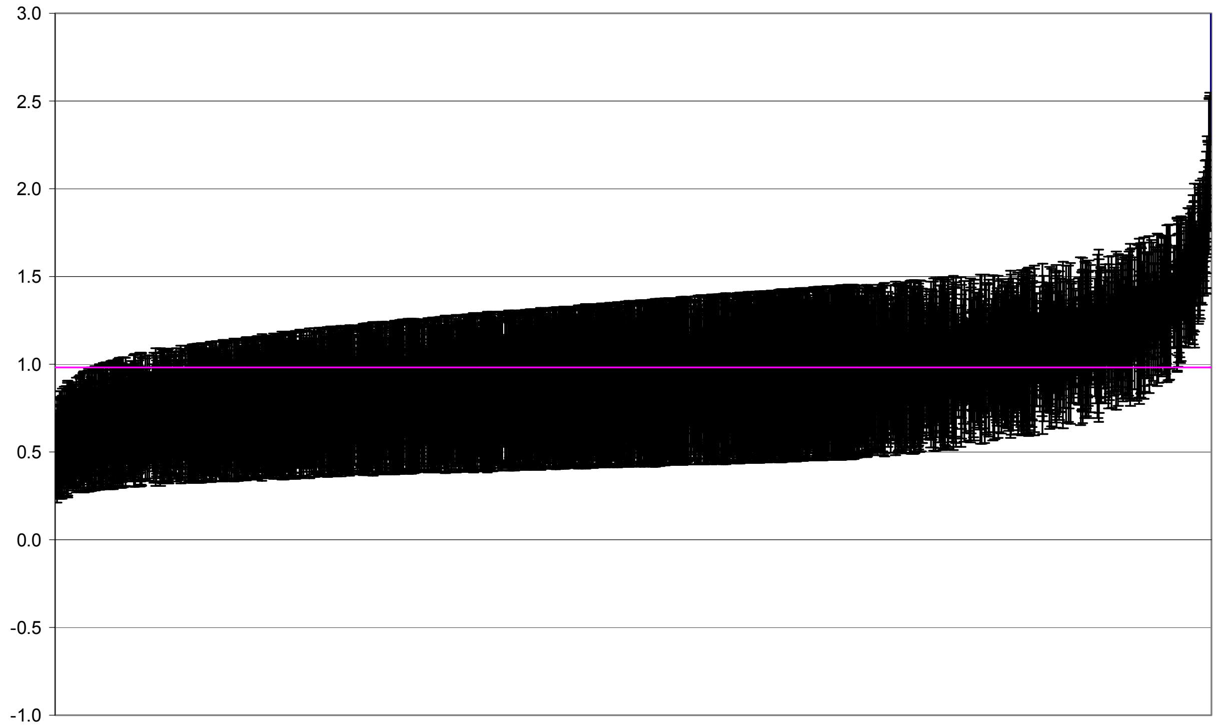

Figure 4-5. Patient Safety for Selected Indicators, Factor Weight

Appendix D. Empirical Methods

D.1. Introduction

The AHRQ Quality Indicator risk-adjustment modules begin with estimating a simple logistic model of a 0/1 outcome variable and a set of patient-level covariates as dependent variables, and using the results to form the predicted outcome for each patient (e.g. P=pr(outcome=1)).

<u>Notation:</u>

Y_{ij} = 0 or 1, outcome for patient j in hospital i.
X_{ij} = covariates (e.g., gender, age, DRG, comorbidity)
P_{ij} = predicted probability from logit of Y on X
= $\exp(X_{ij}\beta)/[1+\exp(X_{ij}\beta)]$
where β is estimated from logit on entire sample.
e_{ij} = $Y_{ij} - P_{ij}$ = logit residual (difference between actual and expected).
n_i = number of patients in sample at hospital i.
α = average outcome in the entire sample[5] (e.g. Ybar).

D.2 Computing the Noise Variance

Estimate the Risk Adjusted Ratio (RAR) and Noise Variance using the Ratio Method[6] of Indirect Standardization for each Hospital:

D.2.1. Estimating RAR:

let $O_i = (1/n_i)\sum(Y_{ij})$ be the observed rate at hospital i
let $E_i = (1/n_i)\sum(P_{ij})$ be the expected rate at hospital i

RAR_i
$= \alpha(O_i/E_i) = \alpha\,[(1/n_i)\sum(Y_{ij})]/\,[(1/n_i)\sum(P_{ij})]$ (where sum is for $j = 1$ to $j = n_i$)
= population rate * observed/expected at hospital i.

D.2.2. Estimating Variance of RAR (SE is the square root):

$Var(RAR_i)$
$= Var[\alpha(O_i/E_i)]$
$= (\alpha/E_i)^2 Var[O_i]$ (since $var(aX) = a^2 var(X)$ for any constant a)
$= (\alpha/E_i)^2 Var[(1/n_i)\sum(Y_{ij})]$ (by the definition of O_i)
$= (\alpha/E_i)^2 (1/n_i)^2 Var[\sum(Y_{ij})]$ (since $var(aX) = a^2 var(X)$ for any constant a)
$= (\alpha/E_i)^2 (1/n_i)^2 [\sum Var(Y_{ij})]$ (since $var(\sum X_i)=\sum var(X_i)$ if Xi are independent)
$= (\alpha/E_i)^2 (1/n_i)^2 \sum [P_{ij}(1-P_{ij})]$ (since Y is 0/1, so $var(Y) = P(1-P)$)

[5] For the AHRQ QI, the sample is the entire reference population consisting of the discharges in the SID for the participating states pooled over three years (2001-2003). Therefore, the "average outcome for the entire sample" is the population rate.
[6] Risk-adjusted rate = (Observed rate / Expected Rate) * Population Rate

D.3. Computing the Composite Variance

1) Setup[7]
   a) Let M be a 1xK vector of observed quality measures (for a given hospital, suppress hospital subscript for convenience), noisy measures of the true underlying 1xK quality vector μ, so that:
      i) $M = \mu + \varepsilon$
      ii) Let the KxK signal variance-covariance be $Var(\mu) = \Omega_\mu$
      iii) Let the KxK noise variance-covariance be $Var(\varepsilon) = \Omega_\varepsilon$
   b) Let $\hat{\mu}$ (1xK) be the posterior (filtered) estimate of μ, so that:
      i) $\mu = \hat{\mu} + \nu$, where the 1xK vector ν represents the prediction error of the posterior estimates, and Var(ν) is the KxK variance-covariance matrix for these posterior estimates.
   c) The goal is to estimate the variance for any weighted average of the posterior estimates. For a given (Kx1) weighting vector (w), this is given by:
      i) $Var(\nu w) = w' Var(\nu) w$
         Thus, we simply need an estimate of Var(ν).

2) Special Case: Filtered estimates are formed in isolation for each measure (univariate) and the estimation error is assumed not correlated across measures (e.g. each measure based on different sample of patients or independent patient outcomes).
   a) Forming each measure in isolation, using superscripts to indicate the measure (k=1,…,K) as above, so:
      i) $\hat{\mu}^k = M^k \hat{\beta}^k = M^k \left[\Omega_\mu^{kk} + \Omega_\varepsilon^{kk}\right]^{-1} \Omega_\mu^{kk}$
      ii) $Var(\nu^k) = \Omega_\mu^{kk} - \Omega_\mu^{kk}\left(\Omega_\mu^{kk} + \Omega_\varepsilon^{kk}\right)^{-1}\Omega_\mu^{kk} = \Omega_\mu^{kk}\left(1 - \hat{\beta}^k\right)$
      iii) Note that in this simple case the filtered estimate is a simple shrinkage estimator and:
         (1) $\hat{\beta}^k$ is the signal ratio of measure k, also is the reliability of the measure, and is the r-squared measuring how much of the variation in the true measure can be explained with the filtered measure.
         (2) The variance of the filtered estimate is simply the signal variance times one minus the signal ratio. Thus, if the signal ratio is zero (no information in the measure), the error in the estimate is equal to the signal variance. But as the signal ratio grows, the error in the estimate shrinks (to zero if there is a signal ratio of 1 – no noise).
   b) The formula for $Var(\nu^k)$ above provides the diagonal elements of $Var(\nu)$ (the full KxK variance-covariance matrix of the filtered estimates). So, get the covariance elements, which are (for j≠k):
      i) $Cov(\nu^j, \nu^k) = E\left[(\mu^j - \hat{\mu}^j)(\mu^k - \hat{\mu}^k)\right]$

[7] For more information on the empirical bayes estimator methods, see the technical appendix in Dimick JB, Staiger DO, Birkmeyer JD. Are Mortality Rates for Different Operations Related?: Implications for Measuring the Quality of Noncardiac Surgery. Medical Care. 44(8):774-778, August 2006; and McClellan M and Staiger D, The Quality of Healthcare Providers, NBER Working Paper #7327, September, 1999 (at http://www.nber.org/papers/w7327).

ii) After some algebra (assuming independent estimation error in the two measures), one gets the following simple expression:

(1) $Cov\left(\nu^{j},\nu^{k}\right)=\Omega_{\mu}^{jk}\left(1-\hat{\beta}^{j}\right)\left(1-\hat{\beta}^{k}\right)$

iii) Note that this is just the signal covariance, times one minus the signal ratio for each of the measures. Thus, if the signal ratio is zero for each measure, the covariance in the estimates is simply the signal covariance. As either measure gets a stronger signal ratio (becomes more precise), the covariance in the estimates shrinks to zero.

iv) Also note that if one measure is missing, then the signal ratio is simply set to zero – the filtered estimate is shrunk all the way back to the (conditional) mean, and the variance and covariance are as defined above.